CATGAAGGTTCCTCCTGTCCTGCTTCTCTTTCTTCTGTCCTCAGTGCGAGCTACTGAGCAACCGCAGGTCG
TCACTGAGCATCCCAGCATGGAGGCAGCCCTGACCGGGCCCAACGCCTCCTCGCACTTCTGGGCCAACTAC
ACTTTCTCTGACTGGCAGAACTTCGTGGGCAGGAGACGTTATGGGGCCGAGTCCCAGAACCCCACGGTGAA
AGCACTGCTCATCGTGGCCTACTCATTCACCATCGTCTTCTCGCTCTTCGGTAATGTCCTGGTCTGTCATG
TCATCTTCAAGAACCAGCGCATGCACTCGGCCACCAGCCTCTTCATTGTCAACCTGGCAGTGGCGGACATC
ATGATCACATTGCTCAACACGCCCTTCACTTTGGTCCGCTTTGTGAACAGCACATGGGTGTTTGGGAAGGG
CATGTGTCATGTCAGTCGCTTTGCTCAGTACTGTTCTCTACATGTCTCAGCACTGACTCTGACAGCTATCG
CAGTGGACCGCCACCAGGTCATCATGCATCCACTGAAGCCTCGGATCTCCATCACCAAGGGTGTCATATAT
ATTGCTGTCATCTGGGTCATGGCTACCTTCTTCTCTCTGCCACATGCCATCTGCCAGAAACTGTTTACCTT
CAAGTACAGTGAGGACATTGTGCGCTCCCTCTGCCTGCCGGACTTCCCGGAGCCAGCTGACCTCTTCTGGA
AGTATCTGGACCTGGCCACCTTCATCCTGCTCTACCTACTTCCACTCTTCATTATCTCAGTGGCCTATGCT
CGTGTGGCCAAGAAGCTGTGGCTCTGTAACACCATTGGCGACGTGACCACAGAGCAGTACCTCGCCCTGCG
ACGCAAGAAGAAGACCACCGTGAAGATGCTGGTGCTTGTGGTAGTCCTCTTTGCCCTCTGCTGGTTCCCTC
TCAACTGCTATGTCCTCCTCTTGTCCAGCAAGGCCATCCACACCAACAATGCCCTCTACTTTGCCTTCCAC
TGGTTTGCCATGAGCAGTACTTGTTATAACCCCTTCATCTACTGCTGGCTCAATGAGAACTTTAGGGTTGA
GCTTAAGGCATTGCTGAGCATGTGCCAAAGGCCACCCAAGCCGCAGGAAGACAGGCTACCCTCCCCAGTTC
CTTCCTTCAGGGTGGCATGGACAGAGAAGAGCCATGGTCGGAGGGCTCCACTACCTAATCACCACTTGCCC
TCTTCCCAGATCCAGTCTGGGAAGACAGATCTGTCATCTGTGGAACCCGTTGTGGCCATGAGTTAGGGAAA
GCTGGAAGTTGGTGGGGGAGGGTTCTTTCCTCTCACAATTGACCAGACACTAACAGAGTTGGAAAGTAACA
CAGAAGCAGTGAGATGCTTGGGTTCCTAGGAACCTGTCCAGCCCCATCTGATTTGCAAACTTTCTAGAAGA
TGCCATGAGGTGGTGTGTGTAGATCTTTGAGCAAGAGCTCTGGAAACCACCTCAGCTTCAACAGAGGCTGG
TCCAGTCAACCACCTCCAATTGTGTAGCATCTGCCACCTTGCCCTTCCTACTGCTGAGCAACCACAGGGGG
ACTTGAGCCATACTATTGGTGGGCCTGCCCCACATGCTCAGAAAAGAACAGGCACAAAGGCTTTCTGAAGT
CATTGGAACAGGAATAATCACACAGCTTCAGTGACCTTGGCTCTATCCATGACCAGACAGGACCCATTTTG
GCTTCTTAAAAACAAAGAGAAATTAGTATTGCCACTTTGAAAAGTTCAGAAAAGTAAAGAAATGAGTTCAG
CCCTCAATTTGTAAAAAAAGGAAAAAAGAAAAAAAAAAGAAAAAGAAAGAAAAAAGCCTGTTAATATGCTG
TAAATTTATCTGTAGCTTTGCCTTCTGTGTGTGTACATTTGTACTTTTAAAATCCTGAACTACACGTGTCC
ATGTAGATTGTAATAATTAGCAAGAAACTGGAATATATCAGAGTATTATTGAATTC (SEQ ID NO:1)

MKVPPVLLLFLLSSVRATEQPQVVTEHPSMEAALTGPNASSHFWANYTFSDWQNFVGRRRYGAESQNPTVK
ALLIVAYSFTIVFSLFGNVLVCHVIFKNQRMHSATSLFIVNLAVADIMITLLNTPFTLVRFVNSTWVFGKG
MCHVSRFAQYCSLHVSALTLTAIAVDRHQVIMHPLKPRISITKGVIYIAVIWVMATFFSLPHAICQKLFTF
KYSEDIVRSLCLPDFPEPADLFWKYLDLATFILLYLLPLFIISVAYARVAKKLWLCNTIGDVTTEQYLALR
RKKKTTVKMLVLVVVLFALCWFPLNCYVLLLSSKAIHTNNALYFAFHWFAMSSTCYNPFIYCWLNENFRVE
LKALLSMCQRPPKPQEDRLPSPVPSFRVAWTEKSHGRRAPLPNHHLPSSQIQSGKTDLSSVEPVVAMS
(SEQ ID NO:2)

Figure 1

Underlined = deleted in targeting construct

Bold = sequence flanking Neo insert in targeting construct

GGGGTGGCAGTCGGCACCATCAGGCTCCCTTGGCGTTTCGGAGTTTTCTCTGTGGTCCCG
ACTCTCCGGAGGATCTCGGTTGTCTCCCAAGTCGGAACCTGGCACGGTCCAGGTTCACTC
GGAGGTCCGGGCTTCCTCTGTGCCCCGTGCCCCTCGCTCCCAGGCTCCCTCTGTGGTGTG
GACTCCTCTAGCCCGGTGCGCTCAGCCCCTCGCACCCAGCCTCCAGGCACAGAGCCCGGC
AGGGAGCTCAGCCCTTGTGCCTAGAGCTGCAGTGGCTGGACATGAAGGTTCCTCCTGT<u>CC</u>
<u>TGCTTCTCTTTCTTCTGTCCTCAGTGCGAGCTACTGAGCAACCGCAGGTCGTCACTGAGC</u>
<u>ATCCCAGCATGGAGGCAGCCCTGACCGGGCCCAACGCCTCCTCGCACTTCTGGGCCAACT</u>
<u>ACACTTTCTCTGACTGGCAGAACTTCGTGGGCAGGAGACGTTATGGGGCCGAGTCCCAGA</u>
<u>ACCCCACGGTGAAAGCACTGCTCATCGTGGCCTACTCATTCACCATCGTCTTCTCGCTCT</u>
<u>TCGGTAATGTCCTGGTCTGTCATGTCATCTTCAAGAACCAGCGCATGCACTCGGCCACCA</u>
<u>GCCTCTTCATTGTCAACCTGGCAGTGGCGGACATCATGATCACATTGCTCAACACGCCCT</u>
<u>TCACTTTGGTCCGCTTTGTGAACAGCACATGGGTGTTTGGGAAGGGCATGTGTCATGTCA</u>
<u>GTCGCTTTGCTCAGTACTGTTCTCTACATGTCT</u>**CAGCACTGACTCTGACAGCTATCGCAG**
TGGACCGCCACCAGGTCATCATGCATCCACTGAAGCCTCGGATCTCCATCACCAAGGGGTG
TCATATATATTGCTGTCATCTGGGTCATGGCTACCTTCTTCTCTCTGCCACATGCCATCT
GCCAGAAACTGTTTACCTTCAAGTACAGTGAGGACATTGTGCGCTCCCTCTGCCTGCCGG
ACTTCCCGGAGCCAGCTGACCTCTTCTGGAAGTATCTGGACCTGGCCACCTTCATCCTGC
TCTACCTACTTCCACTCTTCATTATCTCAGTGGCCTATGCTCGTGTGGCCAAGAAGCTGT
GGCTCTGTAACACCATTGGCGACGTGACCACAGAGCAGTACCTCGCCCTGCGACGCAAGA
AGAAGACCACCGTGAAGATGCTGGTGCTTGTGGTAGTCCTCTTTGCCCTCTGCTGGTTCC
CTCTCAACTGCTATGTCCTCCTCTTGTCCAGCAAGGCCATCCACACCAACAATGCCCTCT
ACTTTGCCTTCCACTGGTTTGCCATGAGCAGTACTTGTTATAACCCCTTCATCTACTGCT
GGCTCAATGAGAACTTTAGGGTTGAGCTTAAGGCATTGCTGAGCATGTGCCAAAGGCCAC
CCAAGCCGCAGGAAGACAGGCTACCCTCCCCAGTTCCTTCCTTCAGGGTGGCATGGACAG
AGAAGAGCCATGGTCGGAGGGCTCCACTACCTAATCACCACTTGCCCTCTTCCCAGATCC
AGTCTGGGAAGACAGATCTGTCATCTGTGGAACCCGTTGTGGCCATGAGTTAGGGAAAGC
TGGAAGTTGGTGGGGGAGGGTTCTTTCCTCTCACAATTGACCAGACACTAACAGAGTTGG
AAAGTAACACAGAAGCAGTGAGATGCTTGGGTTCCTAGGAACCTGTCCAGCCCCATCTGA
TTTGCAAACTTTCTAGAAGATGCCATGAGGTGGTGTGTGTAGATCTTTGAGCAAGAGCTC
TGGAAACCACCTCAGCTTCAACAGAGGCTGGTCCAGTCAACCACCTCCAATTGTGTAGCA
TCTGCCACCTTGCCCTTCCTACTGCTGAGCAACCACAGGGGGACTTGAGCCATACTATTG
GTGGGCCTGCCCCACATGCTCAGAAAAGAACAGGCACAAAGGCTTTCTGAAGTCATTGGA
ACAGGAATAATCACACAGCTTCAGTGACCTTGGCTCTATCCATGACCAGACAGGACCCAT
TTTGGCTTCTTAAAAACAAAGAGAAATTAGTATTGCCACTTTGAAAAGTTCAGAAAAGTA
AAGAAATGAGTTCAGCCCTCAATTTGTAAAAAAAGGAAAAAAGAAAAAAAAAAAGAAAAAG
AAAGAAAAAAGCCTGTTAATATGCTGTAAATTTATCTGTAGCTTTGCCTTCTGTGTGTGT
ACATTTGTACTTTTAAAATCCTGAACTACACGTGTCCATGTAGATTGTAATAATTAGCAA
GAAACTGGAATATATCAGAGTATTATTGAATTC

Figure 2A

Gene Sequence Structure *

299 bp Sequence Deleted 753 bp
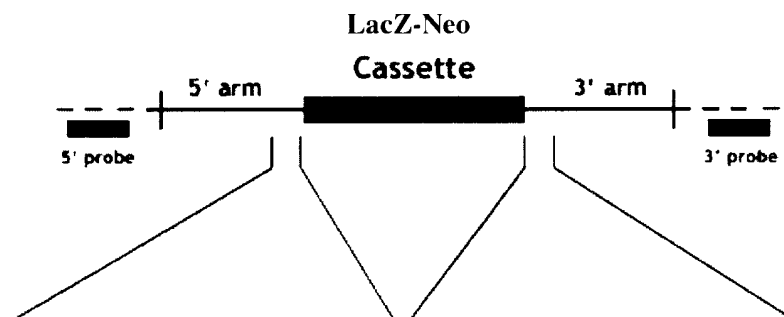
Size of partial cDNA: 2253 bp

Targeting Vector* (genomic sequence)

Construct Number: 463

Arm Length:
5': 2.5 kb
3': 0.6 kb

———— Targeting Vector
— — — — Endogenous Locus

* Not drawn to scale

LacZ-Neo
Cassette

5' arm   3' arm

5' probe   3' probe

5'>CTGGCACGGTCCAGGTTCACT
CGGAGGCCCGGGCTTCCTCTGTGC
CCCGTGCCCCTCGCTCCCTGGCTC
CCTCTGTGGTGTGGACTCCTCTAG
CCCGGTGCGCTCAGCCCCTCGCAC
CCAGCCTCCAGGCACAGAGCCCGG
CAGGGAGCTCAGCCCTTGTGCCTA
GAGCTGCAGTGGCTGGACATGAAG
GTTTCTCCTGT<3'
(SEQ ID NO:3)

5'>CAGCACTGACTCTGACAGCTA
TCGCAGTGGACCGCCACCAGGTGA
GAGCACCTGTCCCCAGCAGCATGC
TCCCATCTCCGTCTATGCCTGGCT
GGCTGGTGGGAATACTGCCACCAC
GGTCTGTAGGGAATACTCTCAGGA
CAGTGACTCATTCAGTCCCGCTGA
CAGCGTGTGTGCTTGCCTCCTTGT
TGATCAATTTG<3'
(SEQ ID NO:4)

Figure 2B